IBM

CHARACTER RECOGNITION

Magnetic and Optical Character Reading

Highly organised societies such as our own depend to a very high degree on efficient administration, and so on accurate and economical paper handling. In our personal affairs we rarely spare a second thought to the convenience of having a bank account and of being able to pay for goods and services by writing a cheque; but without the banking service which has been built up over many centuries, the business of the country would grind to a halt.

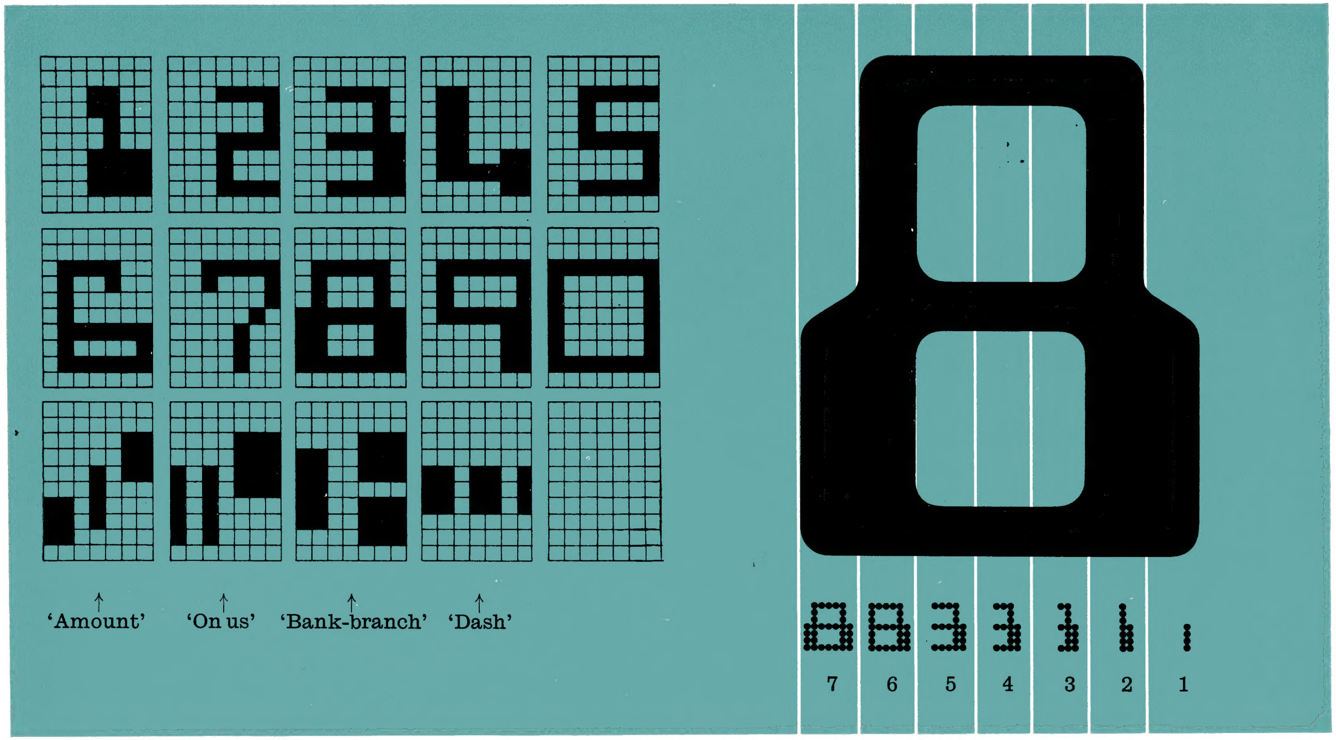
The growth of centuries has been much accelerated in the last few years. In 1950 a total of 321 million articles, representing £94,241 million, were passed through the London Clearing House. In 1960 the number of articles cleared had risen to 438 million, and their cash value to £225,417 million.

When the volume of paper to be handled grows in this way, new methods of coping with it must be investigated. Banks work in close co-operation in order to provide their customers with an efficient service, and the Clearing Banks Electronics Committee was set up to advise on the application of modern techniques to the problems of banking.

The choice of a common machine language was one of the matters discussed by this committee, which has recommended the adoption of the E 13 B typeface printed in magnetic ink along the lower edge of a cheque. Several British banks have already installed or ordered IBM machines which can sort up to 950 documents a minute by this means.

But there are many applications where magnetic ink printing is not practicable, and for these IBM also makes optical character readers, which can read information printed 'in clear' on documents of all sorts. The field of optical character readers is a wide one, covering all data processing which relies for input on information which has been previously recorded with this purpose in mind. In fact, it may reasonably be claimed that this ability to read printed documents for computer input extends the application of automatic data processing to the whole complex of the country's industrial, commercial and social life. Insurance, pension funds, health services, industrial production – all these services and activities depend for their strength on the contributions which individuals make, and the whole effort depends on processing the papers which record each individual transaction.

IBM makes and markets data processing systems which process facts and figures at great speed. It also provides the equipment which turns information into data which can be processed automatically by a computer. This booklet describes two such machines, the IBM 1412 Magnetic Character Reader and the IBM 1418 Optical Character Reader.



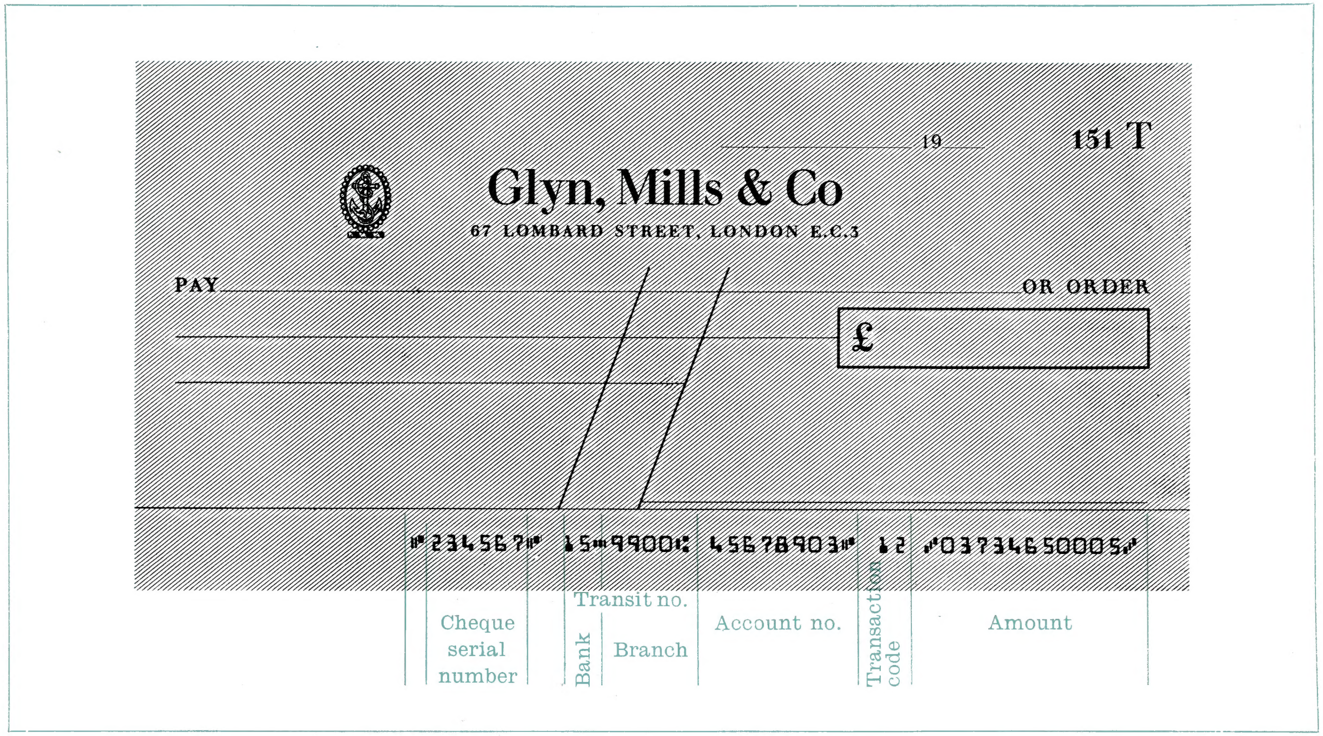
The principle of electronic reading is most simply illustrated by showing how characters printed in magnetic ink are sensed and interpreted.

The multi-channel reading head breaks up the characters into 10 horizontal elements. Seven separate readings are made as the character passes under the reading head, so that each numeral or special character is divided into 70 blocks, or sensing areas.

The presence or absence of magnetic ink in each of these sensing areas is recorded in an electronic register, which therefore builds up an image of the printed character.

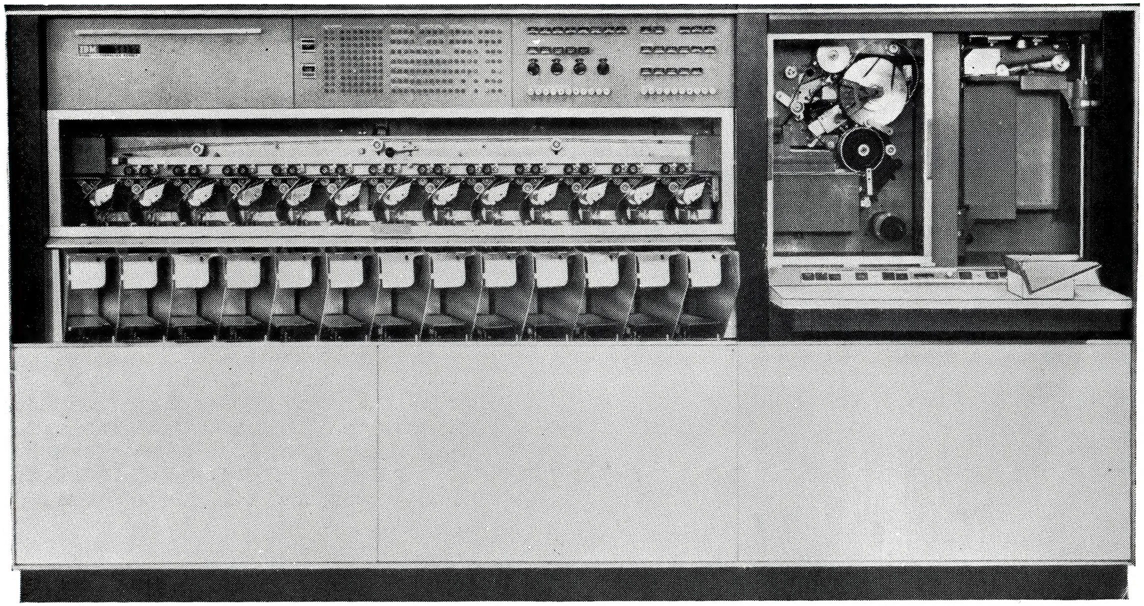
Powerful logical circuits compare this image with the perfect character, and with the thousands of permitted variations. Some variation is inevitable, and the machine's usefulness would be very limited if only perfect reproductions were accepted. It therefore rejects only those images in the electronic register which could be related to more than one of the 14 possible characters.

The characters which make up the E 13 B typeface have been carefully designed after exhaustive analyses of comparative shapes. Each is as different as possible from all the others, yet retains its legibility for the human reader. This style has been adopted by all British and American banks for use in machine reading systems.



The cheques of one bank may be handled, cleared or collected by any of the others, so the Clearing Banks Electronics Committee has decided on a common format to be used by all banks for the code line on cheques. Careful consideration resulted in the format illustrated. The first three fields are printed on the cheques when they are issued to the customer. The last two fields are inscribed on the individual cheque as a by-product of proving or an add/listing operation.

IBM 1412 Magnetic Character Reader-Sorter



Reads and sorts up to 950 documents per minute.

Handles cheques and cards of all sizes within the Clearing Banks Electronics Committee's specifications – length 6 – 8 inches, width 3 – 4 inches.

Provides direct input to an IBM 1401 or 1410 data processing system.

Has two platforms for document feeding without interruption.

Solid-state for compactness, reliability and serviceability.

Sorting on any field

The cheques pass lengthways under the reading head so that all characters along the cheque are scanned each time it is passed through the machine. The field on which sorting is to be performed is selected by depressing one of the five field keys. In order to arrange documents in sequence they must first be sorted on the units digit, then on the tens digit, and so on. There is a further set of keys to select the appropriate digit within the field.

Checking the validity of all characters in all fields

Because the whole code line is scanned each time the cheques are read, validity checks can be performed on all the inscribed data in one run. If the characters have been badly damaged in handling, the cheque will be sorted into the reject pocket for scrutiny; but the accuracy of IBM's unique multi-channel method of character sensing is such that positive identification is possible even though a portion of the character may be missing.

Verifying the number of characters in all fields

Even though every character was valid, a reading could still be wrong if a character were missing entirely. The number of digits in each field is therefore counted, and if this does not correspond with the field length stored in the machine the document will be rejected.

Verifying the account number by means of the self-checking number

Each customer has an account number which gives the bank a numerical identification of his account. To safeguard against any error of transcription the account number ends in a check digit, which is calculated by

The IBM 1412 Magnetic Character Reader-Sorter can simultaneously perform the following six operations:

1. Sort on any digit
2. Check the validity of all characters in all fields
3. Verify the number of characters in all fields
4. Verify the account number by means of the self-checking number
5. Count the number of documents passing through the machine
6. Read data directly into an IBM 1401 or 1410 Data Processing System

In addition, it can do one of the following:

Segregate documents bearing a given account number
Edit documents to ensure that they contain a given transit number
Segregate documents bearing several given transaction code numbers
Accumulate amounts and print totals for batch proving
Check documents for proper sequence

applying a simple formula to all the other digits. The 1412 goes through the same process when reading the account number and independently calculates the check digit. If this does not agree with the printed check digit the document is rejected.

Counting the number of documents passing through the machine

The auxiliary document counter counts the total number of documents in the batch on the first pass through the machine.

Reading data directly into a 1401 or 1410 system

Besides the task of physically sorting the documents, the information contained in them must be processed, and an IBM 1401 data processing system, which can include up to six magnetic tape units and 20 million alphanumerical characters in random access disc storage, offers many advantages for such applications as banking. Two 1412s may be included in the faster and more powerful 1410 system. Since the computer can process the data before instructing the 1412 which pocket to sort the document into, the 1412 will often be used for direct input on the first run, and the documents will be sorted according to the results of complex calculation. Thereafter the 1412 will be used off-line as a sorter.

Segregating documents bearing a given account number

It is often convenient to separate certain documents - *eg* all the dividend warrants issued by a very large company, or all the cheques for a high-activity branch - on the first run through the 1412, without completing the process of sorting. This can be done with the multiple column selector, at the same time as sorting on the units position of a field. This device scans up to six digits and extracts all documents bearing a particular number. All the other documents are sorted normally.

Editing documents to ensure that they contain a given transit number

The multiple column selector can also be used to edit. Instead of selecting documents bearing a given number and sorting the rest, it can sort documents bearing a given number and select the rest. Thus while sorting a batch of cheques received from the clearing bank into their correct sequence, the 1412 could check that they all bear the correct transit number.

Segregating documents bearing several given transaction code numbers

Another form of selection is possible on a single digit. When sorting on a particular digit, sort suppression keys can be used to sort into pockets only those documents which bear certain digits, the remaining documents remaining in sequence in the reject pocket. The incorporation of a code digit in the transaction code can thus be used, for example, to select all credit slips or batch totals from the batch.

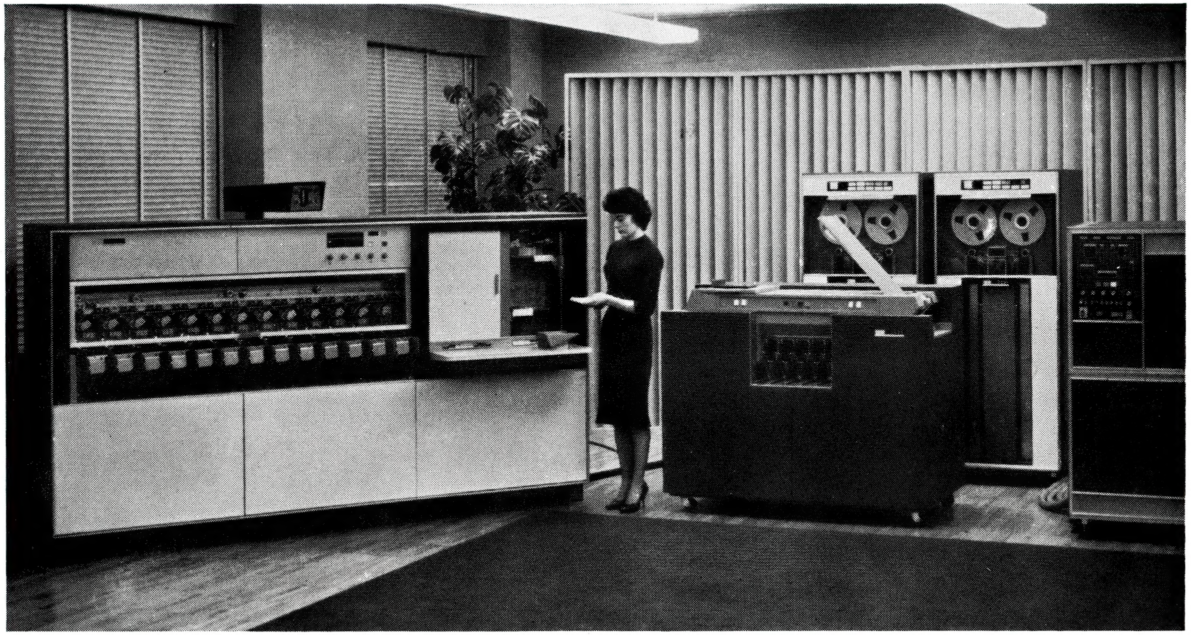
Accumulating amounts and printing totals for batch proving

Batches of cheques which are received accompanied by a listing or a control total can be proved during the first pass through the 1412 by an electronic accumulator. This optional device accumulates amounts from the cheques and prints totals on a paper tape. The speed of reading and sorting is not affected by this process.

Checking documents for proper sequence

After the first sorting pass the electronic accumulator may be used for another purpose, to check the accuracy of the previous sort. Not that the 1412 will have made mistakes in sorting, but it is not impossible that the sequence of the documents may be disturbed while the operator removed them from the pockets and replaced them in the feed. By reading the digit column that was sorted on the previous pass, the electronic accumulator can check for the rare accident.

IBM 1418 Optical Character Reader-Sorter



Reads and sorts up to 400 documents per minute, depending on length*.

Handles cheques and cards of all sizes within the limits – length $5\frac{7}{8}$ – $8\frac{3}{4}$ inches, width $2\frac{3}{4}$ – $3\frac{2}{3}$ inches.

Provides direct input to an IBM 1401 data processing system.

Has two platforms for document feeding without interruption.

Solid-state for compactness, reliability and serviceability.

Data printed with ordinary ink can be read by the 1418 for direct input to an IBM 1401 system, and in order to sort the documents into numerical sequence. The 1418 is capable of reading 480 characters per second, and a second optical reading station may be added so that two lines can be read in one pass through the machine, with a combined input speed to a computer of 960 characters per second. The position of the line to be read is controlled by the operator, by means of the line selection feature.

The 1418 can read numerical data, together with three special characters, printed in IBM 407 typeface. Documents may therefore have been printed on the IBM 407, on the IBM 1403 printer, on an IBM typewriter fitted with the correct typeface, or by plastic plate credit card imprinters with the correct typeface.

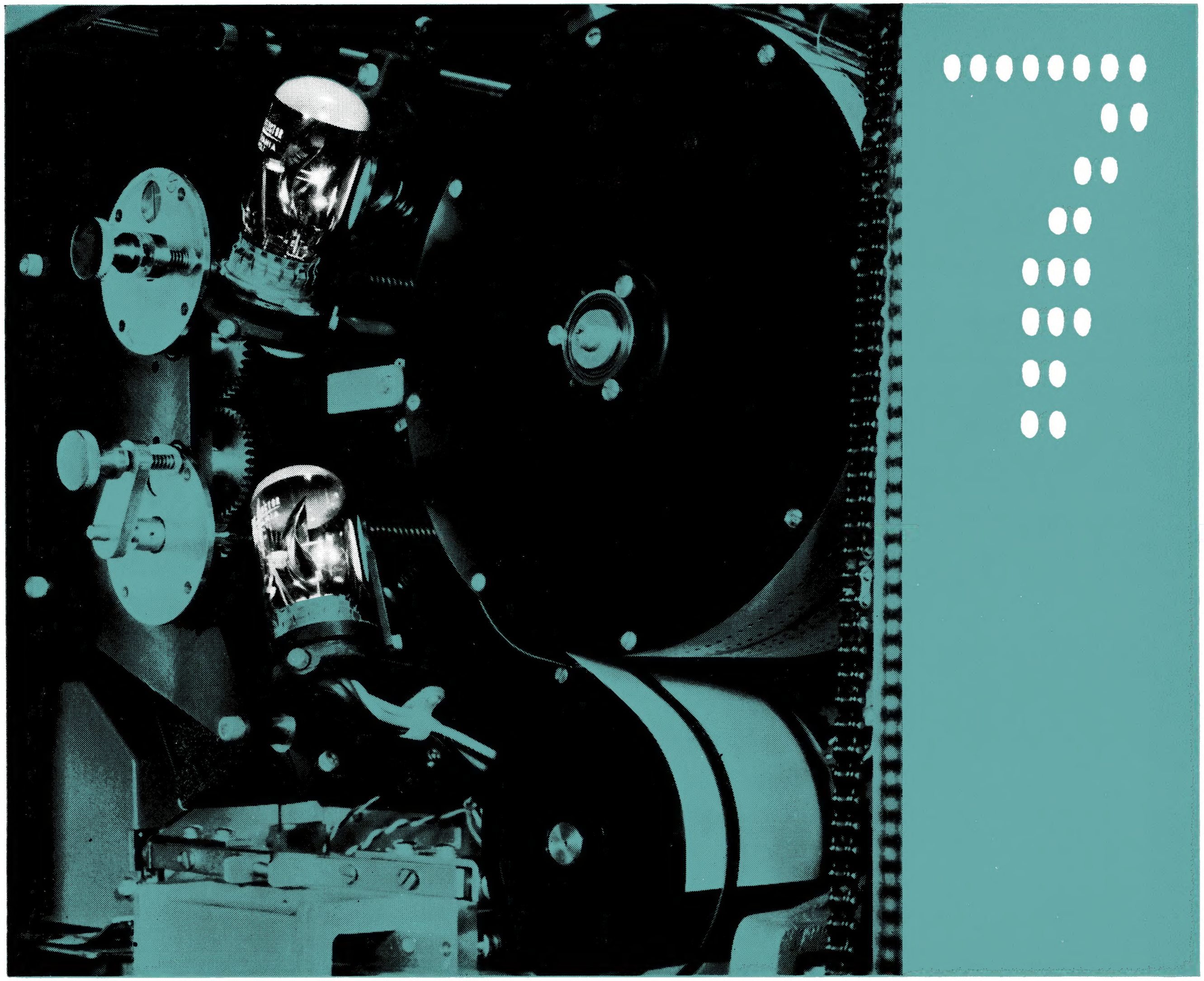
Instead of the second optical reading station, a mark reading station may be added to the 1418. This device enables the 1418 to read information marked on the documents with a pen or pencil, coded in the form of vertical strokes. Their position on the form causes these marks to represent specific characters. The applications of this feature – like those of mark sensing on a reproducing card punch – include meter reading. The meter reader would have a preprinted document showing all the information required except the current meter reading, which he would code on to the document, ready for immediate processing.

Like the 1412, the 1418 may be used both on-line, for direct input to a 1401 system, and off-line for sorting on one digit at a time.

*Documents per minute $= \frac{2880}{L+1.25}$
where L is the length of the document in inches.

The high-resolution method of scanning employed by the IBM 1418 breaks up each character into 170 sensing areas. Such great detail, accompanied by powerful logic, ensures the accuracy of reading, yet keeps the proportion of rejected characters to a minimum.

Lamps flood the face of the document with light, which is reflected through lens assemblies and converted into 'bits' which pass into the electronic register. To ensure proper alignment the document is held in place while it is being read by a series of rollers and a vacuum drum.



Machine-printed documents, such as invoices, can now be used as input to a data processing system when they are returned by customers. As a result, many industries can make use of a 1401–1418 system. For those who do not require numerical sorting of documents, the 1418 is available with only three pockets. Typical applications include:

Insurance
Premium payments
Paid claims

Public utilities
Customer accounting
Meter reading transcription

Retailing
Invoicing and accounts receivable

Hire purchase
Instalment loan accounting

Book and record clubs
Subscription renewal
Invoicing and accounts receivable

Government and local government
Registration and licensing
Tax accounting

Share registration
Dividend cheques

IBM Service

IBM's reputation in more than 90 countries has been built on a tradition going back to the earliest days of punched card data processing machines, a tradition of selling not machines alone, but a comprehensive service – expert sales and technical assistance, educational facilities, preventive maintenance and service bureaux and data centre facilities, as well as the widest choice of the most modern data processing machines in the world.

For further information on how IBM can help you, please get in touch with the branch manager at your nearest branch.

IBM United Kingdom Limited
101 Wigmore Street, London W1
Welbeck 6600

Birmingham
Norfolk House, Smallbrook
Ringway, Birmingham 5
Midland 8921

Bristol
30-34 Baldwin Street, Bristol 1
Bristol 293201

Glasgow
Royal London House
54 West Nile Street, Glasgow C1
Central 8461

Leeds
21 Eastgate, Leeds 2
Leeds 35921

Liverpool
30 Exchange Street East, Liverpool 2
Maritime 1944

London City
41 Moorgate, London EC2
Metropolitan 9901

London West
Yeoman House, Kew Bridge
Brentford, Middlesex
Chiswick 8464

Manchester
Queens House, Queen Street
Manchester 2
Deansgate 8404

Newcastle
Norham House
12 New Bridge Street, Newcastle-on-Tyne 1
Newcastle-on-Tyne 28906

Government
101 Wigmore Street, London W1
Welbeck 6600

Data Centre and Education Centre
58-62 Newman Street, London W1
Langham 6481

Printed in England
UK Form 24–6105